Journal of Machine Learning Research 1 (2000) 1-48 Submitted 4/00; Published 10/00

Accelerated oblique random survival forests

Marina Meilă MMP@STAT.WASHINGTON.EDU
Department of Statistics
University of Washington
Seattle, WA 98195-4322, USA

Michael I. Jordan JORDAN@CS.BERKELEY.EDU
Division of Computer Science and Department of Statistics
University of California
Berkeley, CA 94720-1776, USA

Editor: Kevin Murphy and Bernhard Schölkopf

Abstract

abstract here

Keywords: Bayesian networks, mixture models, Chow-Liu trees

1. Introduction

2. Related work

Several machine learning algorithms can engage with right-censored time-to-event outcomes. In the current study, we consider four classes of learners: random forests, boosting ensembles, regression models, and neural networks.

accelerated oblique random survival forests `aorsf`
original oblique random survival forests `obliqueRSF`
axis-based random survival forests `randomForestSRC & ranger`
axis-based conditional inference forests `party`
gradient boosted decision trees `xgboost`

3. Results

	Performance metric (SD)		Computing time, seconds: Fit model		Computing time, seconds: Predict risk	
	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
actg						
Ranger	0.740 (0.042)	0.053 (0.034)	0.316	0.334	0.120	1.95
Xgboost	0.739 (0.041)	0.044 (0.052)	3.496	3.69	0.028	0.461
aORSF(i=1)	0.735 (0.041)	0.046 (0.050)	0.947	1.00	0.062	1.00
obliqueRSF	0.734 (0.038)	0.049 (0.041)	42.337	44.7	4.027	65.3
aORSF(i=15)	0.733 (0.039)	0.044 (0.048)	2.199	2.32	0.056	0.904

(continued)

	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
Party	0.733 (0.036)	0.048 (0.044)	1.950	2.06	2.795	45.3
aorsf-net	0.729 (0.037)	0.041 (0.052)	29.028	30.6	0.060	0.969
Rfsrc	0.719 (0.040)	0.027 (0.067)	0.288	0.304	0.056	0.908
Coxtime	0.585 (0.127)	-0.004 (0.052)	9.717	10.3	0.458	7.42
breast						
aORSF(i=1)	0.742 (0.046)	0.116 (0.045)	4.732	1.00	0.300	1.00
obliqueRSF	0.739 (0.049)	0.112 (0.042)	2854.152	603.2	6.075	20.3
aORSF(i=15)	0.739 (0.048)	0.117 (0.048)	11.412	2.41	0.289	0.964
Party	0.736 (0.054)	0.102 (0.039)	7.197	1.52	1.862	6.21
Xgboost	0.731 (0.050)	0.094 (0.038)	5.372	1.14	0.274	0.913
aorsf-net	0.730 (0.048)	0.020 (0.086)	611.280	129.2	0.299	0.996
Ranger	0.728 (0.047)	0.087 (0.035)	0.281	0.059	0.171	0.570
Rfsrc	0.688 (0.040)	0.046 (0.043)	0.436	0.092	0.152	0.508
Coxtime	0.659 (0.080)	-0.007 (0.123)	17.179	3.63	1.307	4.36
colon						
Rfsrc	0.816 (0.020)	0.283 (0.029)	1.469	0.470	0.172	0.884
aORSF(i=15)	0.787 (0.021)	0.219 (0.030)	7.471	2.39	0.190	0.975
aORSF(i=1)	0.785 (0.018)	0.219 (0.027)	3.127	1.00	0.195	1.00
aorsf-net	0.784 (0.020)	0.215 (0.029)	159.569	51.0	0.190	0.977
Xgboost	0.783 (0.018)	0.231 (0.025)	8.686	2.78	0.049	0.252
obliqueRSF	0.782 (0.019)	0.219 (0.021)	294.395	94.2	8.161	41.9
Party	0.753 (0.019)	0.185 (0.020)	3.273	1.05	6.168	31.7
Ranger	0.741 (0.021)	0.151 (0.017)	1.030	0.329	0.631	3.24
Coxtime	0.531 (0.052)	-0.003 (0.005)	14.933	4.78	2.590	13.3
flchain						
aORSF(i=1)	0.831 (0.013)	0.334 (0.025)	14.475	1.00	1.238	1.00
aORSF(i=15)	0.831 (0.013)	0.333 (0.024)	29.203	2.02	1.231	0.994
aorsf-net	0.829 (0.013)	0.330 (0.024)	442.726	30.6	1.219	0.985
Xgboost	0.829 (0.012)	0.331 (0.021)	7.123	0.492	0.267	0.216
Ranger	0.828 (0.013)	0.325 (0.022)	100.469	6.94	17.866	14.4
Party	0.827 (0.013)	0.325 (0.024)	33.053	2.28	81.937	66.2
obliqueRSF	0.825 (0.013)	0.325 (0.024)	1800.099	124.4	59.969	48.4
Rfsrc	0.825 (0.012)	0.323 (0.022)	5.793	0.400	0.453	0.366
Coxtime	0.582 (0.109)	-0.015 (0.067)	22.765	1.57	13.661	11.0
gbsg2						
Ranger	0.733 (0.045)	0.118 (0.030)	0.373	0.401	0.124	1.83
obliqueRSF	0.732 (0.036)	0.151 (0.040)	130.378	140.0	2.473	36.4
aORSF(i=15)	0.731 (0.037)	0.146 (0.046)	2.116	2.27	0.061	0.899
Party	0.731 (0.038)	0.146 (0.044)	0.691	0.742	1.008	14.8
aORSF(i=1)	0.730 (0.037)	0.144 (0.049)	0.931	1.00	0.068	1.00
aorsf-net	0.729 (0.036)	0.144 (0.047)	57.048	61.3	0.062	0.910

(continued)

	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
Xgboost	0.728 (0.037)	0.126 (0.037)	2.580	2.77	0.029	0.428
Rfsrc	0.723 (0.035)	0.139 (0.047)	0.618	0.664	0.086	1.27
Coxtime	0.519 (0.058)	-0.057 (0.134)	10.882	11.7	0.680	10.0
guide_it						
aORSF(i=1)	0.725 (0.032)	0.124 (0.038)	1.663	1.00	0.110	1.00
aORSF(i=15)	0.724 (0.034)	0.123 (0.039)	4.896	2.94	0.106	0.963
aorsf-net	0.724 (0.034)	0.119 (0.040)	84.799	51.0	0.110	0.994
obliqueRSF	0.714 (0.037)	0.111 (0.037)	235.729	141.8	3.299	29.9
Ranger	0.713 (0.035)	0.104 (0.032)	0.446	0.268	0.188	1.70
Party	0.709 (0.036)	0.102 (0.032)	1.525	0.917	1.728	15.7
Coxtime	0.699 (0.032)	0.077 (0.053)	12.268	7.38	0.565	5.13
Rfsrc	0.698 (0.033)	0.098 (0.034)	0.846	0.509	0.101	0.915
Xgboost	0.688 (0.032)	0.073 (0.037)	4.948	2.98	0.033	0.295
Mayo Clinic Primary Biliary Cholangitis Data, N = 276						
Xgboost	0.874 (0.011)	0.218 (0.019)	11.715	0.359	0.365	0.123
aORSF(i=15)	0.867 (0.011)	0.221 (0.015)	77.709	2.38	2.874	0.967
aORSF(i=1)	0.867 (0.011)	0.219 (0.016)	32.619	1.00	2.970	1.00
aorsf-net	0.867 (0.011)	0.221 (0.016)	628.839	19.3	3.054	1.03
obliqueRSF	0.865 (0.012)	0.217 (0.015)	2098.616	64.3	270.503	91.1
Party	0.861 (0.013)	0.214 (0.014)	111.579	3.42	460.042	154.9
Ranger	0.859 (0.012)	0.189 (0.013)	171.040	5.24	89.041	30.0
Rfsrc	0.858 (0.012)	0.215 (0.014)	11.646	0.357	1.420	0.478
Coxtime	0.590 (0.105)	0.000 (0.001)	19.733	0.605	50.705	17.1
nafld						
aORSF(i=1)	0.796 (0.077)	0.210 (0.149)	8.429	1.00	0.683	1.00
aORSF(i=15)	0.796 (0.078)	0.210 (0.148)	20.977	2.49	0.668	0.978
aorsf-net	0.793 (0.080)	0.202 (0.160)	270.097	32.0	0.685	1.00
obliqueRSF	0.792 (0.075)	0.187 (0.115)	1025.773	121.7	37.539	55.0
Xgboost	0.789 (0.082)	0.187 (0.180)	6.797	0.806	0.213	0.311
Party	0.787 (0.077)	0.190 (0.131)	24.327	2.89	56.237	82.3
Rfsrc	0.786 (0.085)	0.198 (0.155)	2.467	0.293	0.285	0.417
Ranger	0.785 (0.075)	0.170 (0.119)	19.567	2.32	8.448	12.4
Coxtime	0.641 (0.148)	0.038 (0.211)	18.944	2.25	7.395	10.8
Overall						
aORSF(i=1)	0.926 (0.032)	0.484 (0.078)	0.327	1.00	0.031	1.00
aorsf-net	0.924 (0.031)	0.481 (0.077)	20.545	62.8	0.027	0.861
obliqueRSF	0.924 (0.034)	0.442 (0.066)	165.436	505.8	0.671	21.5
aORSF(i=15)	0.922 (0.031)	0.474 (0.074)	0.711	2.17	0.025	0.786
Party	0.920 (0.041)	0.432 (0.073)	0.276	0.844	0.319	10.2
Xgboost	0.917 (0.035)	0.458 (0.085)	4.339	13.3	0.033	1.04
Rfsrc	0.916 (0.033)	0.444 (0.079)	0.101	0.308	0.033	1.07

(continued)

	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
Ranger	0.903 (0.034)	0.349 (0.041)	0.035	0.107	0.042	1.33
Coxtime	0.562 (0.201)	-0.012 (0.018)	12.376	37.8	0.322	10.3
peakV02						
aorsf-net	0.721 (0.022)	0.112 (0.034)	174.090	37.6	0.320	1.04
obliqueRSF	0.720 (0.022)	0.119 (0.030)	614.000	132.8	10.231	33.3
aORSF(i=15)	0.720 (0.020)	0.115 (0.031)	10.718	2.32	0.302	0.983
aORSF(i=1)	0.720 (0.020)	0.113 (0.032)	4.625	1.00	0.308	1.00
Party	0.718 (0.023)	0.117 (0.029)	5.315	1.15	9.747	31.7
Ranger	0.714 (0.022)	0.103 (0.023)	2.211	0.478	1.358	4.42
Rfsrc	0.714 (0.019)	0.111 (0.026)	1.718	0.372	0.193	0.627
Xgboost	0.714 (0.018)	0.102 (0.028)	5.134	1.11	0.047	0.152
Coxtime	0.701 (0.028)	0.097 (0.032)	20.249	4.38	3.560	11.6
Rotterdam Breast Cancer Data, N = 2,982						
aorsf-net	0.724 (0.012)	0.151 (0.018)	260.864	43.4	0.377	0.954
aORSF(i=1)	0.723 (0.013)	0.150 (0.019)	6.007	1.00	0.395	1.00
aORSF(i=15)	0.723 (0.013)	0.150 (0.019)	13.286	2.21	0.390	0.985
Xgboost	0.722 (0.014)	0.143 (0.017)	3.979	0.662	0.089	0.225
Ranger	0.721 (0.013)	0.143 (0.014)	1.881	0.313	1.819	4.60
obliqueRSF	0.720 (0.013)	0.149 (0.017)	912.920	152.0	18.351	46.4
Party	0.719 (0.012)	0.147 (0.016)	6.186	1.03	15.125	38.3
Rfsrc	0.716 (0.013)	0.137 (0.020)	2.762	0.460	0.261	0.661
Coxtime	0.502 (0.049)	-0.047 (0.154)	18.276	3.04	5.857	14.8
sprint-acm						
Xgboost	0.798 (0.014)	0.205 (0.026)	13.680	0.447	0.199	0.087
aORSF(i=15)	0.795 (0.013)	0.199 (0.020)	87.852	2.87	2.271	0.992
aORSF(i=1)	0.795 (0.013)	0.199 (0.022)	30.597	1.00	2.288	1.00
aorsf-net	0.795 (0.013)	0.192 (0.023)	934.189	30.5	2.298	1.00
obliqueRSF	0.794 (0.013)	0.194 (0.019)	4216.468	137.8	81.441	35.6
Party	0.792 (0.012)	0.187 (0.017)	93.194	3.05	130.262	56.9
Rfsrc	0.791 (0.011)	0.192 (0.020)	6.913	0.226	0.496	0.217
Ranger	0.784 (0.013)	0.175 (0.016)	9.100	0.297	8.720	3.81
Coxtime	0.768 (0.087)	0.177 (0.054)	46.094	1.51	15.082	6.59
sprint-cvd						
aorsf-net	0.822 (0.018)	0.136 (0.016)	532.742	24.2	1.615	1.01
Party	0.821 (0.015)	0.121 (0.009)	92.038	4.18	126.234	79.0
aORSF(i=15)	0.821 (0.018)	0.131 (0.011)	59.100	2.68	1.591	0.995
aORSF(i=1)	0.821 (0.018)	0.131 (0.011)	22.035	1.00	1.599	1.00
obliqueRSF	0.821 (0.016)	0.128 (0.009)	1569.905	71.2	92.265	57.7
Xgboost	0.819 (0.016)	0.135 (0.019)	10.625	0.482	0.099	0.062
Rfsrc	0.817 (0.014)	0.130 (0.013)	3.344	0.152	0.451	0.282
Ranger	0.812 (0.017)	0.113 (0.008)	5.724	0.260	6.390	4.00

(continued)

	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
Coxtime	0.787 (0.021)	0.098 (0.024)	20.202	0.917	13.321	8.33
time-to-million						
Xgboost	0.933 (0.018)	0.553 (0.055)	13.028	9.08	0.034	0.420
Coxtime	0.931 (0.016)	0.563 (0.057)	20.527	14.3	0.751	9.27
aORSF(i=15)	0.931 (0.017)	0.534 (0.038)	4.789	3.34	0.076	0.941
aorsf-net	0.930 (0.017)	0.535 (0.039)	78.002	54.4	0.077	0.952
aORSF(i=1)	0.929 (0.018)	0.530 (0.040)	1.435	1.00	0.081	1.00
Rfsrc	0.925 (0.018)	0.527 (0.041)	0.822	0.573	0.066	0.812
Party	0.912 (0.024)	0.466 (0.043)	0.617	0.430	0.905	11.2
obliqueRSF	0.909 (0.024)	0.337 (0.031)	239.423	166.9	2.436	30.0
Ranger	0.904 (0.024)	0.436 (0.042)	0.076	0.053	0.125	1.54
vdv						
Xgboost	0.778 (0.096)	-0.065 (0.330)	4.236	1.55	1.617	2.78
Rfsrc	0.773 (0.118)	0.047 (0.189)	0.185	0.068	0.304	0.523
aORSF(i=15)	0.772 (0.091)	0.053 (0.126)	2.841	1.04	0.552	0.948
aORSF(i=1)	0.772 (0.082)	0.046 (0.119)	2.732	1.00	0.582	1.00
Ranger	0.765 (0.075)	0.024 (0.119)	0.496	0.181	0.099	0.171
obliqueRSF	0.764 (0.076)	0.064 (0.131)	120.857	44.2	2.945	5.06
Party	0.761 (0.088)	0.053 (0.119)	7.889	2.89	5.362	9.21
aorsf-net	0.750 (0.100)	0.055 (0.188)	21.757	7.96	0.552	0.948
Coxtime	0.678 (0.134)	-0.259 (0.417)	29.478	10.8	1.899	3.26
veteran						
aORSF(i=15)	0.846 (0.053)	0.291 (0.072)	0.352	1.90	0.012	0.656
aORSF(i=1)	0.843 (0.053)	0.289 (0.071)	0.186	1.00	0.018	1.00
aorsf-net	0.840 (0.055)	0.281 (0.077)	15.971	86.1	0.012	0.657
Ranger	0.836 (0.072)	0.175 (0.057)	0.019	0.104	0.028	1.56
obliqueRSF	0.835 (0.051)	0.192 (0.051)	91.890	495.3	0.235	13.2
Rfsrc	0.816 (0.064)	0.248 (0.079)	0.057	0.310	0.023	1.27
Party	0.816 (0.061)	0.208 (0.059)	0.117	0.632	0.067	3.75
Xgboost	0.780 (0.069)	0.151 (0.085)	3.009	16.2	0.026	1.47
Coxtime	0.523 (0.077)	-0.038 (0.055)	9.481	51.1	0.159	8.97

Acknowledgments

We would like to acknowledge support for this project from the National Science Foundation (NSF grant IIS-9988642) and the Multidisciplinary Research Program of the Department of Defense (MURI N00014-00-1-0637).

Appendix A.

In this appendix we prove the following theorem from Section 6.2:
Theorem *Let u, v, w be discrete variables such that v, w do not co-occur with u (i.e., $u \neq 0 \Rightarrow v = w = 0$ in a given dataset $\mathcal{D}$). Let N_{v0}, N_{w0} be the number of data points for which $v = 0, w = 0$ respectively, and let I_{uv}, I_{uw} be the respective empirical mutual information values based on the sample $\mathcal{D}$. Then*

$$N_{v0} \; > \; N_{w0} \;\; \Rightarrow \;\; I_{uv} \; \leq \; I_{uw}$$

with equality only if u is identically 0. ■
Proof. We use the notation:

$$P_v(i) \; = \; \frac{N_v^i}{N}, \quad i \neq 0; \quad P_{v0} \; \equiv \; P_v(0) \; = \; 1 - \sum_{i \neq 0} P_v(i).$$

These values represent the (empirical) probabilities of v taking value $i \neq 0$ and 0 respectively. Entropies will be denoted by H. We aim to show that $\frac{\partial I_{uv}}{\partial P_{v0}} < 0$....

Remainder omitted in this sample. See http://www.jmlr.org/papers/ for full paper.